Learning with Multitask Adversaries using Weakly Labelled Data for Semantic Segmentation in Retinal Images

Oindrila Saha OINDRILA_SAHA13@IITKGP.AC.IN
Rachana Sathish RACHANA.SATHISH@IITKGP.AC.IN
Debdoot Sheet DEBDOOT@EE.IITKGP.AC.IN
Indian Institute of Technology Kharagpur

Abstract

A prime challenge in building data driven inference models is the unavailability of statistically significant amount of labelled data. Since datasets are typically designed for a specific purpose, and accordingly are weakly labelled for only a single class instead of being exhaustively annotated. Despite there being multiple datasets which cumulatively represents a large corpus, their weak labelling poses challenge for direct use. In case of retinal images, they have inspired development of data driven learning based algorithms for segmenting anatomical landmarks like vessels and optic disc as well as pathologies like microaneurysms, hemorrhages, hard exudates and soft exudates. The aspiration is to learn to segment all such classes using only a single fully convolutional neural network (FCN), while the challenge being that there is no single training dataset with all classes annotated. We solve this problem by training a single network using separate weakly labelled datasets. Essentially we use an adversarial learning approach in addition to the classically employed objective of distortion loss minimization for semantic segmentation using FCN, where the objectives of discriminators are to learn to (a) predict which of the classes are actually present in the input fundus image, and (b) distinguish between manual annotations vs. segmented results for each of the classes. The first discriminator works to enforce the network to segment those classes which are present in the fundus image although may not have been annotated i.e. all retinal images have vessels while pathology datasets may not have annotated them in the dataset. The second discriminator contributes to making the segmentation result as realistic as possible. We experimentally demonstrate using weakly labelled datasets of DRIVE containing only annotations of vessels and IDRiD containing annotations for lesions and optic disc. Our method using a single FCN achieves competitive results over prior art for either vessel or optic disk or pathology segmentation on these datasets.

Keywords: Adversarial learning, convolutional neural networks, multitask learning, semantic segmentation, retinal image analysis.

1. Introduction

The precise segmentation of retinal anatomies and pathologies serves as an important tool for diagnosis and evaluation of various metabolic and ophthalmic disorders such as diabetes, hypertension, glaucoma and choroidal neovascularization, etc. (Bowling, 2015). Experts can analyze changes in vascular morphology by segmenting the retinal vessels. Vessel and optic disk segmentation followed by localization of fovea region by Haddouche et al. (2010) is the first step for the quantitative analysis of retinal images which is helpful for diabetic retinopathy (DR) screening and diagnosis (Teng et al., 2002). DR is the leading cause of blindness in the working-age population. Screening for DR and monitoring disease progression, especially in the early asymptomatic stages, is effective

for preventing visual loss and reducing costs for health systems (Nentwich and Ulbig, 2015). The most common signs of DR are red lesions symptomatic of microaneurysms, hemorrhages and bright lesions symptomatic of exudates. However, it is tedious and time consuming to segment retinal diseases or anatomies manually, especially in fundus images. Thus ability to automatically and reliably segment all structures is a long-standing problem since decades.

One of the major challenges in this task is the absence of a single dataset which contains exhaustive pixel level semantic annotation of all parts of the retina. However this task of creating such a dataset being a highly taxing job, we propose to employ multiple readily available datasets which have reliable annotations for some of the classes, with no necessity of any single dataset having all given classes annotated and there also not being common classes annotated across different datasets. This paper proposes a single model that can learn from separate datasets to predict all classes in a given retinal image. We address this problem by utilizing multiple adversaries for addressing it as a multitask approach. The convolutional neural network for semantic segmentation tries to generate segmentation maps which visually resemble the ground truth. Two discriminators are used; one for distinguishing between manual vs. segmented maps and the other for determining which of the various classes are present in the fundus image. The first discriminator contributes to making the segmentation result as realistic as possible, thus learning to capture the finer details. The second discriminator works to enforce the model to learn to segment for all classes which are present and are not just limited to those which are annotated in a particular image. Results show the superior performance of the proposed method over several existing methods of vessel segmentation in terms of sensitivity, specificity and classification accuracy. The method reduces over-segmentation i.e. the presence of false positives as compared to previous state-of-the-art. Comparison with the results from the leaderboard in the Diabetic Retinopathy Segmentation and grading challenge [1] show that proposed method achieves better area under Precision vs Sensitivity curve for the different lesions and surpasses Jaccard Index scores for Optic Disk segmentation.

The paper is organized as follows. The existing methods for vessel and lesion segmentation are analyzed in Section 2. A thorough description of the overall proposed adversarial learning method and network architectures is presented in Section 3. Section 4 presents details of the experiments, setup and evaluation criteria. In Section 5, the results are presented and compared to existing methods. Section 6 concludes the paper.

2. Related Work

The existing segmentation methods can be divided into supervised or unsupervised categories according to whether the manual labeled ground truths are used or not.

Azzopardi and Petkov (2013) proposed to use a matched filtering method which convolves a 2-D kernel with the retinal image where the matched filter response indicates the presence of the feature. Salazar-Gonzalez et al. (2014) first carried out a preprocessing for the image by adaptive histogram equalization and robust distance transform, and then segmented retinal vessels with graph cut. Other methods include the vessel profile analysis (Wang et al., 2007), active contour based segmentation (Al-Diri et al., 2009), line detection (Ricci and Perfetti, 2007) and cluster based models (Emary et al., 2014). In case of diabetic retinopathy the first solutions were trained to detect lesions using manual segmentation for supervised classification (Mookiah et al., 2013). Ali et al. (2013) uses a statistical atlas for exudate segmentation, while Inoue et al. (2013) uses Eigen value analysis and

1. https://idrid.grand-challenge.org/Home/

Hessian matrix for microaneurysm segmentation. Most of these models employed a supervised approach for the specific segmentation task and employed curated datasets suitable only for a given task in hand, weakly labelling only a few of all the available classes in the images.

Supervised learning generally requires annotated data in order to build a predictive model. These methods can be regarded as a pixel-level binary classification problem. Each pixel belongs to vessel or non-vessel. There are two parts to segmentation using supervised approaches. One is an extractor to extract the feature vectors of pixels; the other one is a classifier to map extracted vectors to the corresponding labels for each pixel. A number of feature extractors like the Gabor filters (Akram et al., 2014), the Gaussian filter (Zhang et al., 2010) have been used previously. Various classifiers such as k-NN classifier (Staal et al., 2004), support vector machine (SVM) (Gandhi and Dhanasekaran, 2013), artificial neural networks (ANN) (García et al., 2009), AdaBoost (Lupascu et al., 2010) etc, have been proposed to deal with the task.

With the advent of deep learning, there have been numerous studies that investigated the vessel segmentation problems using convolutional neural networks (CNNs). Liskowski and Krawiec (2016) proposed a supervised segmentation architecture that used a deep neural network learned with a large training dataset which was preprocessed via global contrast normalization, zero-phase whitening, geometric transformations and gamma corrections. Tan et al. (2017) used a CNN for segmentation of haemorrhages, exudates and microaneurysms. Fu et al. (2016) regarded the segmentation as a boundary detection problem and they combined the CNN and conditional random field (CRF) layers into an integrated deep network to achieve their goal.

The effectiveness of the method proposed by Luc et al. (2016) and Ganin et al. (2016) for semantic segmentation tasks gives rise to the motivation for using a adversarial approach. The next section describes in detail the proposed multitask adversarial learning method for retinal anatomy and pathology segmentation while learning the single task from multiple weakly labelled datasets.

3. Methodology

In the proposed approach, the challenge of segmentation is formulated as a task of pixel-wise classification of the retinal images using a fully-convolutional neural network. An adversarial approach is used for training of the segmentation network as shown in Figure 1. Given an image $\mathscr{I}$ of size $3 \times M \times N$ the problem of segmentation can be formulated as a pixel-wise classification task where each pixel is assigned a label $\mathscr{L} \in \{ l_1, l_2, \ldots, l_C \}$ such that an output tensor $\mathbf{O}$ of size $C \times M \times N$ is generated, where C is the number of classes. Each channel of $\mathbf{O}$ corresponds to one of the classes: *background, retinal vessels, optic disk, microaneurysms, hemorrhages, hard exudates and soft exudates*. The background is excluded for calculation of segmentation loss and for input to the discriminators, so as to prevent inconsistency in the form of the model learning optic disk as a part of background for DRIVE (Staal et al., 2004) images and retinal vessels as background for IDRiD[2] images. The input to *Discriminator 1* is the retinal image concatenated with the channels of the output segmentation map without the background channel. These channels are randomly shuffled by a *ChannelShuffler()*. It provides a one-hot vector $\hat{\mathbf{n}}_1$ of length $C-1$ in the output which determines which of the channels were present in the input retinal image. In case of a healthy retinal image which consists of only retinal vessels and optic disk without any pathology, the target vector would be $\mathbf{n}_1 = \{1,1,0,0,0,0\}$, while if in a given image the only pathology present is hemorrhage, then $\mathbf{n}_1 = \{1,1,0,1,0,0\}$.

2. http://dx.doi.org/10.21227/H25W98

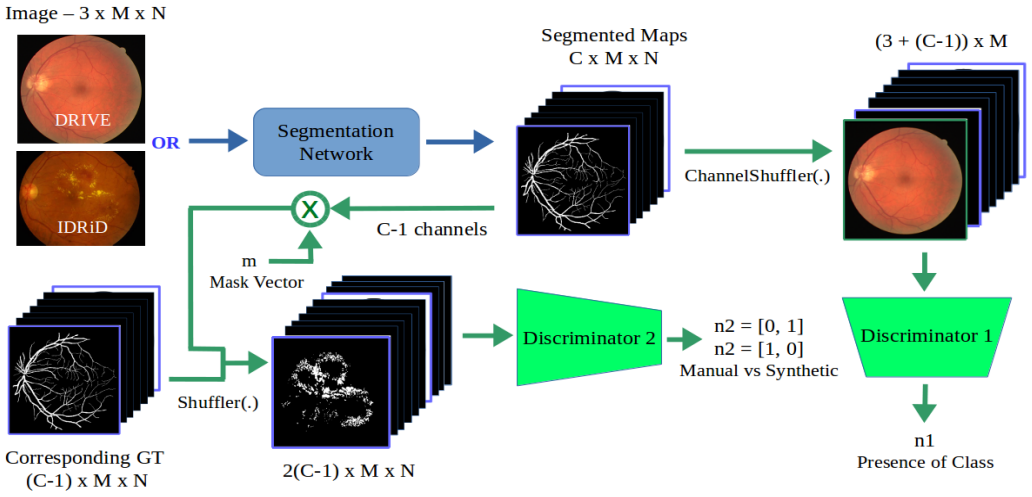


Figure 1: The proposed methodology: Segmentation network outputs a C channel tensor. Discriminator 1 is fed the shuffled segmented maps concatenated with the original image. The segmented maps are masked with vector $\mathbf{m}$ and then concatenated with the groundtruth maps in random order and fed to Discriminator 2.

The task of *Discriminator 2* is to distinguish between manual and segmented maps. The segmented maps are masked with a one-hot vector $\mathbf{m} = \{m_1, m_2, \ldots, m_C\}$, where m_c is 1 if the class is annotated and 0 otherwise. This masking is done so that when the ground truth of a class is not present i.e. blank, the output segmented channel of that class is also masked out to avoid uncertainty in loss measures. Thus, *Discriminator 2* will focus only on differentiating between manual vs. segmented maps for the classes that are annotated in a particular image. The manual and segmented maps are concatenated in random order with probability p from an uniform distribution $p \sim U(0,1)$ and fed in one go. *Discriminator 2* provides as a output a one-hot vector $\hat{\mathbf{n}}_2$ of length 2 for this task. Discriminators are trained using Binary Cross Entropy loss function given as

$$L_{D_i}(\hat{\mathbf{n}}_i, \mathbf{n}_i) = -\frac{1}{C_{D_i}} \sum_{j=1}^{C_{D_i}} (n_{i,j} \log(\hat{n}_{i,j}) + (1 - n_{i,j}) \log(1 - \hat{n}_{i,j})) \tag{1}$$

where $\mathbf{n}_i = \{n_{i,j}\}$, $\hat{\mathbf{n}}_i = \{\hat{n}_{i,j}\}$ are the target and output for the discriminator D_i respectively for $i \in 1,2$ and C_{D_i} denotes the length of the output vector of D_i. *Discriminator 1* corresponds to D_1 with $C_{D_1} = 6$ and *Discriminator 2* corresponds to D_2 with $C_{D_2} = 2$.

The basic segmentation loss is also calculated using Binary Cross Entropy loss. During training of the segmentation network a weighted sum of the segmentation loss and discriminator loss is used.

$$L_{\text{Seg-BCE}}(\mathbf{O}, \mathbf{T}) = -\frac{1}{MN} \sum_{k=1}^{C-1} \sum_{i=1}^{M} \sum_{j=1}^{N} (T_{i,j,k} \log(O_{i,j,k}) + (1 - T_{i,j,k}) \log(1 - O_{i,j,k})) \tag{2}$$

$$L_{Seg}(\mathbf{O}, \mathbf{T}) = \lambda L_{Seg-BCE}(\mathbf{O}, \mathbf{T}) + L_{D_1}(\hat{\mathbf{n}}_\mathbf{1}, \mathbf{n}_\mathbf{1}) + L_{D_2}(\hat{\mathbf{n}}_\mathbf{2}, \mathbf{n}_\mathbf{2}) \tag{3}$$

where the trade-off coefficient λ balances two objective functions. $\mathbf{T} = \{T_{i,j,k}\}$, $\mathbf{O}$ are the target and output for the segmentation network respectively and C denotes the number of classes.

3.1. Dataset Normalization

Retinal images have irregularities in illumination and noise, thus enhancements are performed on the fundus images. On every input image I, contrast limited adaptive histogram equalization is applied to enhance the contrast and also ensure that the images of both the datasets become more similar to each other, for the model to be able to predict classes which are not annotated. A 3×3 median filter is then applied to reduce the noise in background of the image. Finally the intensity values are scaled to $[0, 1]$ to obtain the preprocessed image $\mathscr{I}$ to be operated upon.

3.2. Network Structure

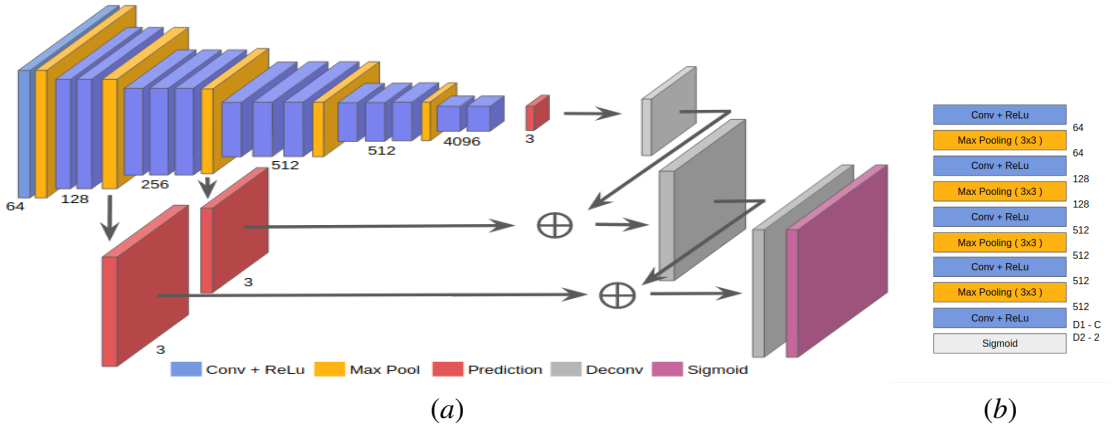
Inspired by the fully convolutional networks (Long et al., 2015) a CNN architecture with skip connections is used as the segmentation network. The main path follows the typical architecture of the VGG 16 network (Simonyan and Zisserman, 2015). It consists of the repeated application of 3×3 convolutions with padding for same size, each followed by a rectified linear unit and a 2×2 max pooling operation with 2 pixels stride for downsampling to reduce the amount of parameters and computation. At each downsampling step the number of feature channels are doubled. In order to concatenate different feature maps through the skip connection, an upsampling of the feature map followed by a 2×2 deconvolution (Long et al., 2015) is used, which is initialized with billinear interpolation filters. This skip-connection is crucial so as to propagate context information to higher resolution layers for more precise segmentation. After a concatenation, each followed by a rectified linear unit, a dropout (Srivastava et al., 2014) layer with probability 0.5 is added for regularization to reduce overfitting. Finally, a sigmoid operation is applied to calculate probability for each class to obtain the output segmentation map $\mathbf{O}$. Fig.2 illustrates the segmentation network. More deconvolution layers were not added as they increased computation with no improvement in performance.

The same architecture is used for both the discriminator networks. It consists of five convolution layers with filter size 4×4, each followed by *LeakyRelu()*, Batch normalization layers (Ioffe and Szegedy, 2015). The first four convolution layers are followed by 3×3 max pooling operations. Finally, a sigmoid layer is used to scale the output to probabilities i.e. in the range $[0, 1]$.

4. Experiments

4.1. Datasets

The DRIVE dataset (Staal et al., 2004) is used for training and testing for the retinal vessels which contains 40 images. The IDRiD dataset (Porwal et al., 2018) contains 81 retinal fundus images with annotations of the optic disk, microaneurysms, haemorrhages, hard exudates and soft exudates. Both datasets are already divided into two standard splits for training and testing. The IDRiD images of size 4288×2848 px were downsampled to 880×584 px while preserving the aspect ratio. DRIVE images were zero padded to a size of 880×584 px.



(*a*) (*b*)

Figure 2: (a) The segmentation network architecture; (b) Basic architecture of both the discriminators: the last convolution layer has output of $C-1$ channels for Discriminator 1 (D1) and 2 channels for Discriminator 2 (D2)

4.2. Training Setup

As data available for training a deep network is less, to overcome this limitation various augmentations are used to increase the number of retinal images for gaining better generalization and preventing overfitting. Images are randomly flipped by probability p generated using a uniform distribution or rotated by an angle θ chosen from an uniform distribution $[-180, 180]$ degrees.

The segmentation network uses the initial layers of VGG network (Simonyan and Zisserman, 2015) and those layers are initialized with the weights of a pretrained VGG model. As the dataset size is small, this initialization improves the results by a considerable margin. The vanilla FCN segmentation model is considered as a baseline. Experiments are conducted where both vanilla segmentation network and the adversarial approach are trained and compared resulting in proposed network achieving higher IOU score.

Adam optimizer (Kingma and Ba, 2014) with learning rate 10^{-4} and $\beta = 0.9$ is used for training all the three networks. The trade-off coefficient λ in (3) is set to 10. In every epoch the discriminators are trained first and then these losses summed with the segmentation loss is backpropagated through the segmentation network. Early stopping of the training based on the validation loss is adopted to prevent overfitting.

4.3. Evaluation

Performance evaluation of segmentation is carried out by considering : sensitivity (SE), specificity (SP), accuracy (ACC), F-score (F1) and Area under the curve (AUC) for vessels, for comparison to previous methods. The AUC is calculated as the area under the SE versus (1-SP) plot. For lesions, the area under precision (PPV) vs sensitivity (SE) curve is computed, while for optic disk the Jaccard index (IOU) is measured so as to compare with the reported results of the Diabetic Retinopathy Segmentation challenge. The code for trying out inference has been released at https://github.com/oindrilasaha/multitask-retinal-segmentation.

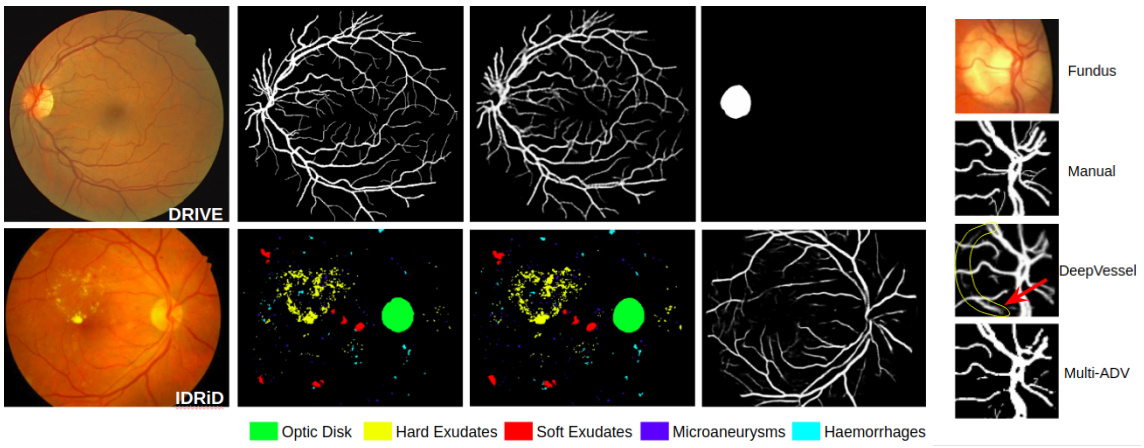


Figure 3: (a) Qualitative results: First column shows the retinal image, second column shows the manual annotation used as ground truth, third column shows the results obtained by the proposed method for the annotated classes in original dataset, and fourth column shows prediction of class not annotated but present in the image; (b) Comparison with DeepVessel (Fu et al., 2016)

5. Results

The proposed method is compared with the existing methods for vessel segmentation in Table 1. SE, SP, F1 and ACC are reported as the maximum value over thresholds. The FCN models trained without discriminator, show inferior performance with respect to Multi-ADV suggesting that the adversarial framework improves quality of segmentation. Our method outperforms existing vessel segmentation methods in literature. Quantitatively, for lesion and optic disc segmentation, the proposed method is compared to the best performing teams from the Diabetic Retinopathy challenge leaderboard[3] using the metrics used in the challenge for ranking in Table 2. Fig. 3(*a*) illustrates the qualitative results of the Multi-ADV model on DRIVE and IDRiD fundus images. It also shows how the method predicts classes which are not annotated for either dataset. It predicts optic disk for DRIVE images and retinal vessels for IDRiD images too. Fig. 3(*b*) compares the result of DeepVessel (Fu et al., 2016) to our method. DeepVessel incorrectly predicts the optic disk boundary as a retinal vessel while it was claimed that their method eliminates incorrect over-segmentation. Our method not only prevents over-segmentation but has also detected more finer vessels than DeepVessel.

6. Conclusion

In this paper, a multiadversary based fully convolutional neural network is proposed for retinal anatomy and pathology segmentation using weakly labelled fundus images. A FCN with skip connections is used as the segmentor network. The skip connections propagate context information to higher resolution layers for more precise segmentation. Two discriminators are used: the first

3. https://idrid.grand-challenge.org/Leaderboard/

Table 1: Comparison with existing models for Vessel Segmentation

Method	SE	SP	ACC	F1	AUC
Azzopardi and Petkov (2013)	0.7655	0.9704	0.9442	-	0.9614
Liskowski and Krawiec (2016)	0.7811	0.9807	0.9535	-	0.9790
DeepVessel (Fu et al., 2016)	0.7603	-	0.9523	-	-
Orlando et al. (2017)	0.7897	0.9684	0.9454	0.7857	0.9506
Alom et al. (2018)	0.7792	0.9813	0.9556	-	0.9784
Maninis et al. (2016)	-	-	-	**0.8220**	-
FCN	0.7731	0.9724	0.9467	0.7612	0.9588
Multi-ADV	**0.7906**	**0.9839**	**0.9641**	0.7925	**0.9812**

Table 2: Comparison with IDRiD leaderboard for Optic Disk and Lesions

Class	Metric	Existing	Proposed	ACC - Proposed
Optic Disk	Jaccard Index (IOU)	0.9338	**0.9644**	0.9829
Microaneurysms	Area under PPV vs SE	0.5017	**0.5504**	0.9034
Haemorrhages	Area under PPV vs SE	0.6804	**0.7338**	0.9655
Soft Exudates	Area under PPV vs SE	0.6995	**0.7118**	0.9618
Hard Exudates	Area under PPV vs SE	**0.8850**	0.8698	0.9771

discriminator contributes to making the segmentation result as realistic as possible while the second discriminator works to enforce segmentation of classes not annotated for a particular image. Results indicates that proposed approach outperforms other state-of-the-arts. The model also predicts classes that were originally not annotated in a given dataset by learning from other dataset.

References

M Usman Akram, Shehzad Khalid, Anam Tariq, Shoab A Khan, and Farooque Azam. Detection and classification of retinal lesions for grading of diabetic retinopathy. *Comp. Biol., Med.*, 45: 161–171, 2014.

Bashir Al-Diri, Andrew Hunter, and David Steel. An active contour model for segmenting and measuring retinal vessels. *IEEE Trans. Med. imaging*, 28(9):1488–1497, 2009.

Sharib Ali, Désiré Sidibé, Kedir M Adal, Luca Giancardo, Edward Chaum, Thomas P Karnowski, and Fabrice Mériaudeau. Statistical atlas based exudate segmentation. *Comp. Med. Imaging, Graph.*, 37(5-6):358–368, 2013.

Md Zahangir Alom, Mahmudul Hasan, Chris Yakopcic, Tarek M Taha, and Vijayan K Asari. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv preprint arXiv:1802.06955*, 2018.

George Azzopardi and Nicolai Petkov. Automatic detection of vascular bifurcations in segmented retinal images using trainable cosfire filters. *Pattern Recognition Letters*, 34(8):922–933, 2013.

Brad Bowling. *Kanski's Clinical Ophthalmology E-Book: A Systematic Approach*. Elsevier Health Sciences, 2015.

Eid Emary, Hossam M Zawbaa, Aboul Ella Hassanien, Gerald Schaefer, and Ahmad Taher Azar. Retinal vessel segmentation based on possibilistic fuzzy c-means clustering optimised with cuckoo search. In *Proc. Int. Jt. Conf. Neur. Networks*, pages 1792–1796, 2014.

Huazhu Fu, Yanwu Xu, Stephen Lin, Damon Wing Kee Wong, and Jiang Liu. Deepvessel: Retinal vessel segmentation via deep learning and conditional random field. In *Proc. Int. Conf. Med. Image Comput., Comp.-Assist. Interv.*, pages 132–139, 2016.

Mahendran Gandhi and R Dhanasekaran. Diagnosis of diabetic retinopathy using morphological process and svm classifier. In *Proc. Int. COnf. Comm., Signal Process.*, pages 873–877, 2013.

Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *J. Mach. Learn. Res.*, 17(1):2096–2030, 2016.

María García, Clara I Sánchez, María I López, Daniel Abásolo, and Roberto Hornero. Neural network based detection of hard exudates in retinal images. *Comp. Methods, Programs, Biomed.*, 93(1):9–19, 2009.

A Haddouche, Mouloud Adel, Monique Rasigni, J Conrath, and Salah Bourennane. Detection of the foveal avascular zone on retinal angiograms using markov random fields. *Digital Signal Processing*, 20(1):149–154, 2010.

Tsuyoshi Inoue, Yuji Hatanaka, Susumu Okumura, Chisako Muramatsu, and Hiroshi Fujita. Automated microaneurysm detection method based on eigenvalue analysis using hessian matrix in retinal fundus images. In *Proc. An. Conf. IEEE Engg., Med., Biol. Soc.*, pages 5873–5876, 2013.

Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *Proc. Int. Conf. Machine Learn.*, pages 448–456, 2015.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014.

Paweł Liskowski and Krzysztof Krawiec. Segmenting retinal blood vessels with deep neural networks. *IEEE Trans. Medi. Imaging*, 35(11):2369–2380, 2016.

Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proc. IEEE/CVF Conf. Comp. Vis., Patt. Recog.*, pages 3431–3440, 2015.

Pauline Luc, Camille Couprie, Soumith Chintala, and Jakob Verbeek. Semantic segmentation using adversarial networks. *arXiv preprint arXiv:1611.08408*, 2016.

Carmen Alina Lupascu, Domenico Tegolo, and Emanuele Trucco. Fabc: retinal vessel segmentation using adaboost. *IEEE Trans. Inf. Tech. Biomed.*, 14(5):1267–1274, 2010.

Kevis-Kokitsi Maninis, Jordi Pont-Tuset, Pablo Arbeláez, and Luc Van Gool. Deep retinal image understanding. In *Proc. Int. Conf. Med. Image Comput., Comp.-Assist. Interv.*, pages 140–148, 2016.

Muthu Rama Krishnan Mookiah, U. Rajendra Acharya, Chua Kuang Chua, Choo Min Lim, E. Y. K. Ng, and Augustinus Laude. Computer-aided diagnosis of diabetic retinopathy: A review. *Comput. Biol. Med.*, 43(12):2136–2155, Dec. 2013.

Martin M Nentwich and Michael W Ulbig. Diabetic retinopathy-ocular complications of diabetes mellitus. *World J. Diabetes*, 6(3):489, 2015.

Jan Odstrcilik, Radim Kolar, Attila Budai, Joachim Hornegger, Jiří Jan, Jiri Gazarek, Tomas Kubena, Pavel Cernosek, Ondrej Svoboda, and Elli Angelopoulou. Retinal vessel segmentation by improved matched filtering: Evaluation on a new high-resolution fundus image database. *IET Image Process.*, 7:373–383, 06 2013.

Jose Orlando, Elena Prokofyeva, and Matthew Blaschko. A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images. *IEEE Trans. Biomed. Engg.*, 64(1):16–27, Jan 2017.

Prasanna Porwal, Samiksha Pachade, Ravi Kamble, Manesh Kokare, Girish Deshmukh, Vivek Sahasrabuddhe, and Fabrice Meriaudeau. Indian Diabetic Retinopathy Image Dataset (IDRiD). IEEE Dataport, 2018.

Elisa Ricci and Renzo Perfetti. Retinal blood vessel segmentation using line operators and support vector classification. *IEEE Trans. Medi. Imaging*, 26(10):1357–1365, 2007.

Ana G Salazar-Gonzalez, Djibril Kaba, Yongmin Li, and Xiaohui Liu. Segmentation of the blood vessels and optic disk in retinal images. *IEEE J. Biomed. Health Inf.*, 18(6):1874–1886, 2014.

Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *Proc. Int. Conf. Learn. Rep.*, 2015.

Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15(1): 1929–1958, 2014.

Joes Staal, Michael D Abràmoff, Meindert Niemeijer, Max A Viergever, and Bram Van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imaging*, 23 (4):501–509, 2004.

Jen Hong Tan, Hamido Fujita, Sobha Sivaprasad, Sulatha V Bhandary, A Krishna Rao, Kuang Chua Chua, and U Rajendra Acharya. Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network. *Information Sciences*, 420:66–76, 2017.

Thomas Teng, Martin Lefley, and D Claremont. Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy. *Med., Biol., Engg., Comp.*, 40(1):2–13, 2002.

Li Wang, Abhir Bhalerao, and Roland Wilson. Analysis of retinal vasculature using a multiresolution hermite model. *IEEE Trans. Med. Imaging*, 26(2):137–152, 2007.

Bob Zhang, Lin Zhang, Lei Zhang, and Fakhri Karray. Retinal vessel extraction by matched filter with first-order derivative of gaussian. *Comp. Biol., Med.*, 40(4):438–445, 2010.

Appendix

A few experimental results are demonstrated on other datasets. Inference only was performed on these images using our existing model without re-training it with any images from these datasets.

Inference on Vessel dataset

HRF dataset [4] was used for inference for vessels class. The results are presented in Table 3. It is worthwhile to note that despite of the network not having had seen HRF images during training, it performs almost equivocal with slightly higher sensitivity and accuracy as compared to the competitive prior art for the task of vessel segmentation.

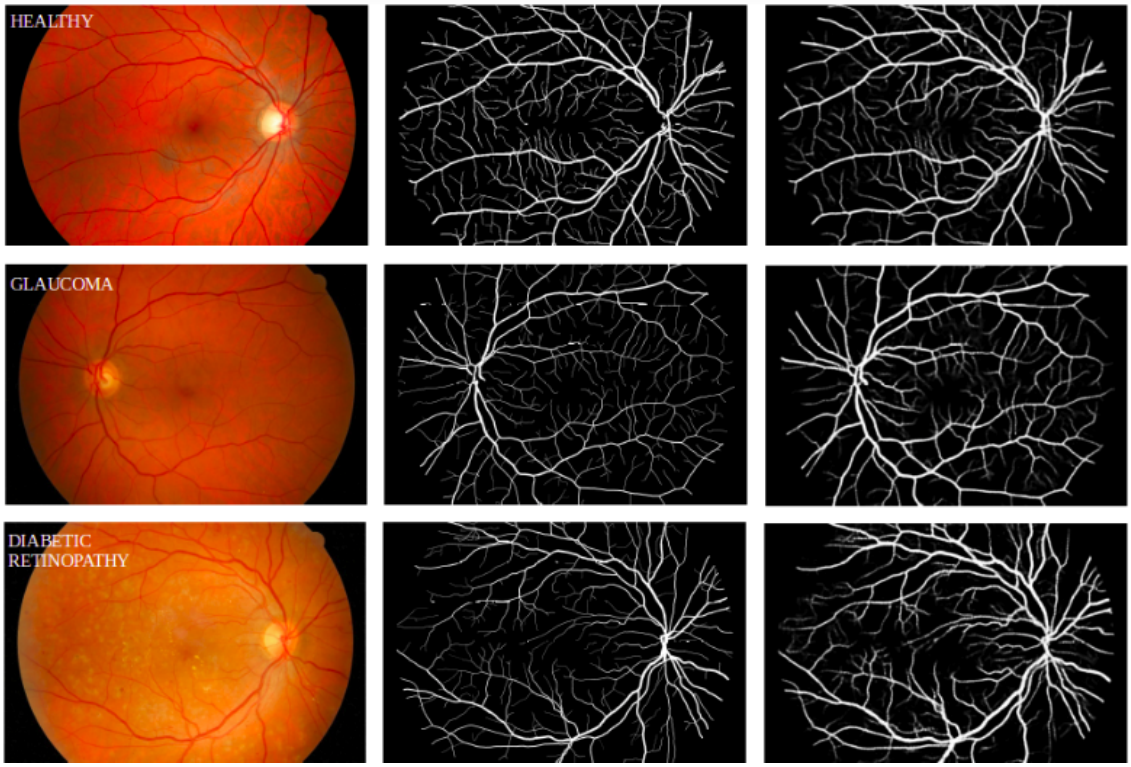


Figure 4: Vessel Segmentation on HRF dataset. Left: Fundus Image, center: Ground Truth, right: Prediction

4. https://www5.cs.fau.de/research/dat/fundus-images/

Table 3: Performance evaluation on HRF dataset

Method	SE	SP	ACC	F1
Proposed	0.7891	0.9642	0.9610	0.7552
Odstrcilik et al. (2013)	0.7741	0.9669	0.9493	-

Inference on pathology datasets

Similar to the approach in for vessel segmentation above, only inferencing is employed on e-optha [5] for exudates and microaneurysms (Table 4), and REFUGE [6] for optic disk (Table 5). For the REFUGE dataset, the results were compared with the challenge leaderboard [7] results.

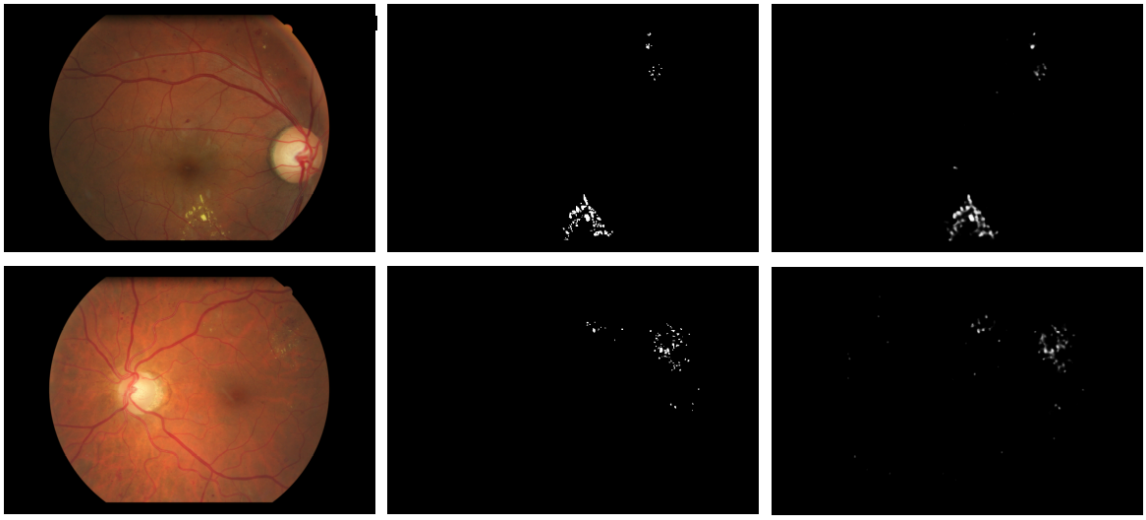


Figure 5: Hard Exudates Segmentation on eoptha-EX dataset. Left: Fundus Image, center: Ground Truth, right: Prediction

Table 4: Performance evaluation on eoptha dataset

Dataset	Area under Precision Recall
eoptha-EX	0.8235
eoptha-MA	0.2500

5. http://www.adcis.net/es/Descargar-Software-Base-De-Datos-Terceros/E-Ophtha.html
6. https://refuge.grand-challenge.org
7. https://refuge.grand-challenge.org/leaderboard

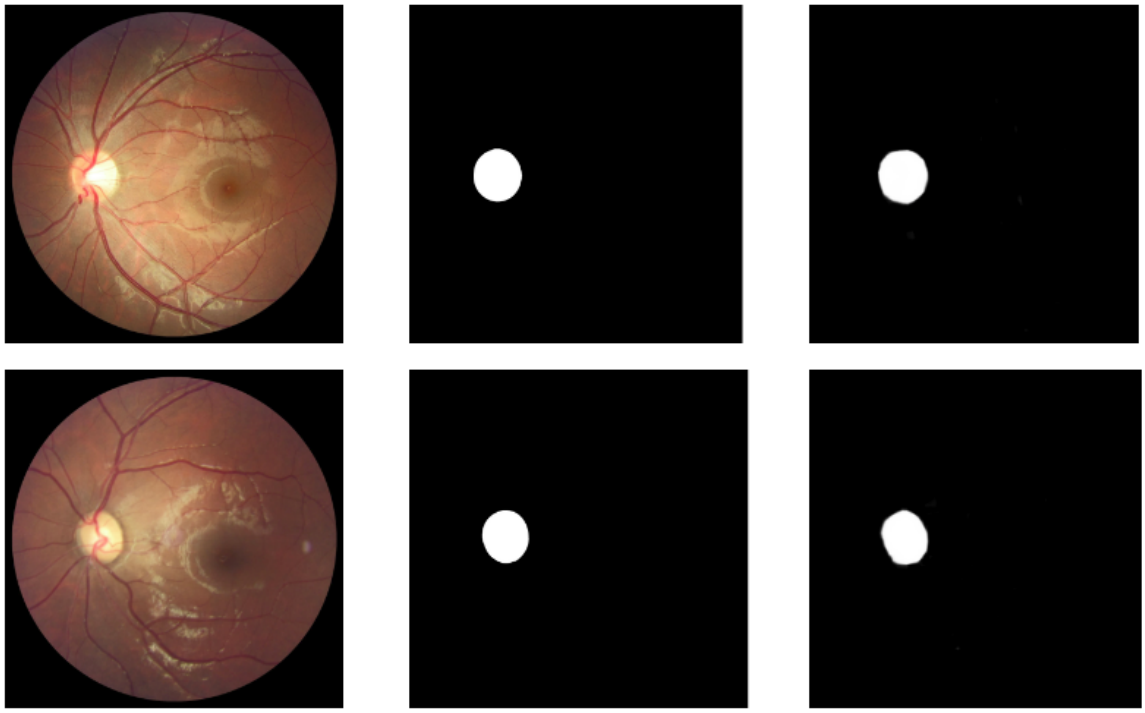


Figure 6: Optic Disk Segmentation on REFUGE dataset. Left: Fundus Image, center: Ground Truth, right: Prediction

Table 5: Performance evaluation on REFUGE dataset

Method	**F1**	**Jaccard**
Proposed	0.9286	0.9202
REFUGE leaderboard	0.958	-